

DOI:10.19651/j.cnki.emt.2211202

改进型 CenterNet 的高铁无砟轨道板 表面裂缝检测算法*

吴铭权 罗晖 李琛彪 李佳敏 蔡联明
(华东交通大学信息工程学院 南昌 330013)

摘要:针对传统的高铁无砟轨道板表面裂缝检测方法存在检测精度低、速度慢的问题,提出一种改进型 CenterNet 的高铁无砟轨道板表面裂缝检测算法。该算法在编解码网络之间加入空洞空间金字塔池化模块(ASPP),以此扩大特征图的感受野,充分提取不同尺度的上下文信息;然后在特征提取网络中加入多光谱通道注意力模块(MCA),使网络可以更好学习每个通道的权重,捕获图像丰富的输入特征信息;最后使用 α -IoU 损失函数来提高边界框预测的准确度。实验结果表明,本算法平均检测精度(mAP)达到 84.12%,相比传统算法平均检测精度提升了 3.37%,对于轨道板表面裂缝具有较好的检测效果。

关键词:裂缝检测;CenterNet;空洞空间金字塔池化;多光谱通道注意力; α -IoU

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Improved CenterNet algorithm for detecting surface cracks in ballastless track slabs of high-speed rail

Wu Mingquan Luo Hui Li Chenbiao Li Jiamin Cai Lianming
(School of Information Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Aiming at the problems of low detection accuracy and slow speed of the traditional method for detecting surface cracks on ballastless track slabs of high-speed railways, an improved CenterNet-based algorithm for detecting surface cracks on track slabs is proposed. The algorithm adds atrous space pyramid pooling module (ASPP) between the codec network as a way to expand the perceptual field of the feature map and fully extract the contextual information at different scales. Then adds a multi-spectral channel attention module (MCA) to the feature extraction network so that the network can better learn the weights of each channel and capture the image rich input feature information. Finally, the α -IoU loss function is used to improve the accuracy of bounding box prediction. The experimental results show that the mean average precision(mAP) of the proposed algorithm reaches 84.12%, which is 3.37% higher than that of the traditional algorithm, and it has a good detection effect on the surface cracks of the track plate.

Keywords: crack detection;CenterNet;atrous space pyramid pooling;multi-spectral channel attention; α -IoU

0 引言

近年来,随着高铁运营里程的增加,轨道板服役时间的延长,在温度应力、列车荷载、风化等多方面因素综合作用下,轨道板会产生裂缝、破损等病害,而轨道板的表面裂缝是最常见的一类病害^[1-2]。轨道板表面裂缝如未能及时发现,会对无砟轨道的结构承载能力、稳定性以及长期服役耐久性造成一定的影响。因此,为了保障高铁线路的安全运

营,实现对轨道板表面裂缝的准确、快速、智能化检测已成为高铁发展亟待解决的关键问题。

传统方法是利用天窗时间通过人工巡检来完成检测轨道板裂缝的任务,然而人工检测方法存在检测精度低及安全隐患等问题。近年来,利用智能化方法来实现轨道板表面裂缝检测成为了一个新的研究方向。Zhang 等^[3]针对传统裂缝检测耗时长、精度低等问题,提出了一种改进的 YOLOv3 的算法,在精度和速度上具有更好的检测效果和

收稿日期:2022-08-28

* 基金项目:国家自然科学基金(62262021)、江西省教育厅科学技术研究重点项目(GJJ200603)资助

鲁棒性;李文举等^[4]采用一种结合投票机制和卷积神经网络的改进方法,有效减少轨道板裂缝检测漏检和误检现象;Ye 等^[5]提出了一种结合扩张卷积的快速检测网络模型 STCNet I,并利用分水岭算法对检测到的轨道板裂缝进行分割,进一步提高了裂缝的检测速度和准确率。

随着深度学习在目标检测领域内进展迅速,运用深度学习的方法来进行轨道板表面裂缝检测是目前研究的热点,也是未来热门研究方向。目前,通过深度学习来实现目标检测的算法主要有两类:一类以 Faster R-CNN^[6-7]、Mask R-CNN^[8]等为代表的双阶段检测算法,它的流程是利用 RPN 生成预选框,再对预选框进行回归分类。另外一类以 YOLO^[9-11]系列、RetinaNet^[12]、CenterNet^[13]等为代表的一阶段检测算法,该类算法无需生成预选框,直接获得检测结果。一阶段检测算法在精度上虽然比双阶段检测算法略低,但是在速度上有着更大的优势,是目前工业应用中最广泛的目标检测算法。

作为一阶段检测算法的 CenterNet 算法,不仅在检测速度上有优势,在检测精度上也表现出色。因此本文提出了一种改进型 CenterNet 的高铁无砟轨道板表面裂缝检测算法,建立轨道板裂缝数据集并对算法进行实验验证分析。

1 CenterNet 算法模型介绍

CenterNet 算法是 Zhou 等^[13]在 2019 年提出的一种无锚框算法,它由编解码网络和检测网络组成,通过关键点热力图来找到中心点,进而根据中心点的图像特征回归得到目标的其他特征,如目标尺寸、方向、姿态信息等,无需预先设置锚点,大大减少了网络参数和计算量。编码器对输入图像进行高级语义特征信息提取,然后通过解码器生成高级语义的特征信息,提升输出特征图分辨率。再利用高斯核将输入图像转换为特征热力图输入检测模块,通过热图表征检测目标信息。检测模块由 3 个预测分支组成,分别为目标中心点预测分支 heatmap、偏移量预测分支 center offset、目标尺寸预测分支 boxes size。该算法选用 ResNet50^[14]为主干网络,其框架如图 1 所示。

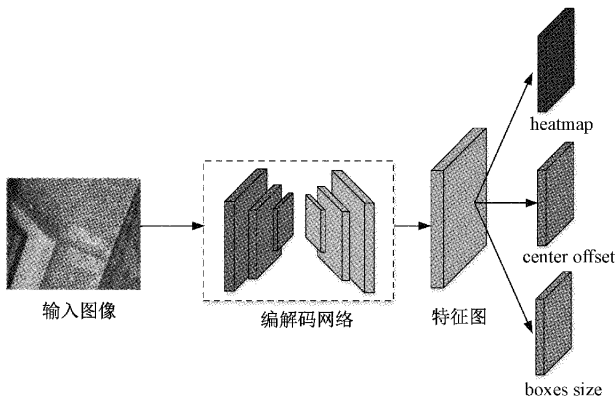


图 1 CenterNet 模型框架

2 改进型 CenterNet 的目标检测算法

在高铁无砟轨道板表面裂缝检测过程中传统检测方法存在检测精度低、速度慢的问题。因此,本文提出的改进型 CenterNet 算法模型,通过引入更优的网络模块和边界框损失函数来提高目标检测能力。

2.1 空洞空间金字塔池化模块

在轨道板表面裂缝检测过程中,极易受到光线阴暗、拍摄角度、复杂场景等因素的干扰,造成拍摄到的部分图片关键信息缺失,进而影响裂缝检测效果。而空洞空间金字塔池化模块(atrous spatial pyramid pooling, ASPP)^[15]较好的解决了这个问题。ASPP 模块通过在空间金字塔池化(spatial pyramid pooling, SPP)^[16]并联多个不同扩张率的空洞卷积,在不丢失图像分辨率的同时扩大特征图的感受野,并获得多个不同尺寸的感受野,从而实现提取图像的不同尺度上下文信息的目的。

本文结合 ASPP 模块改进 CenterNet 算法设计多尺度感受野模块,利用不同扩张率的空洞卷积提取图像的不同尺度上下文信息,其中将 ASPP 三个 3×3 空洞卷积的扩张率分别设定为 6、12、18。如图 2 所示,ASPP 将获得的输入特征图输出到以下 5 个分支中,第 1 个分支是一个 1×1 的卷积,目的是保持原始感受野不变;第 2、3、4 分支 3×3 空洞卷积扩张率依次为 6、12、18,目的是扩大特征图的感受野,从而获取图像的不同尺度上下文信息;第 5 个分支将输入特征做全局平均池化获取全局特征,然后将 5 个不同尺度的特征在通道维度上连接在一起后送入 1×1 卷积核进行融合输出,最后利用双线性插值对低维特征进行上采样操作。本文模型在增加较少计算量的同时,有效识别不同尺度的裂缝特征,降低了裂缝的漏检率。

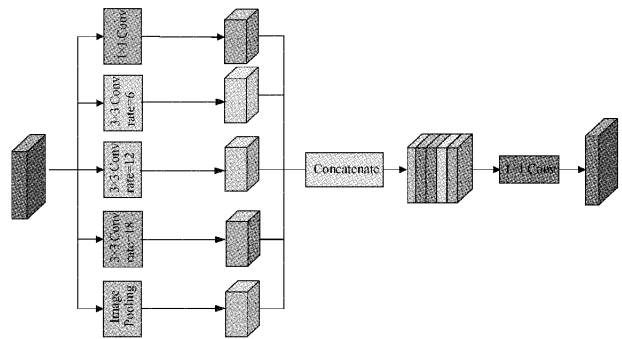


图 2 空洞空间金字塔池化模块结构

2.2 多光谱通道注意力模块

注意力机制本质上是定位到感兴趣的区域同时抑制不相关的背景区域,使得模型聚焦于重要的关键信息,忽略次要信息。尤其是通道注意力机制,在改善深度卷积神经网络(convolutional neural network, CNN)的性能方面有巨大的潜力。然而,大多数传统的通道注意力机制都是通过全局平均池化(global average pooling, GAP)计算每个通

道的权重,如 SEnet^[17],ECAnet^[18],但经过 GAP 得到的输入特征缺乏特征多样性,不利于捕获图像丰富的输入特征信息。本实验选择的多光谱通道注意力模块(multi-spectral channel attention, MCA)^[19]从通道的角度证明了 GAP 是二维离散余弦变换(discrete cosine transform, DCT)的最低频率分量。多光谱通道注意力模块的二维 DCT 定义如下:

$$f_{h,w}^{2d} = \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} X_{i,j}^{2d} \cos\left(\frac{\pi h}{H}\left(i + \frac{1}{2}\right)\right) \cos\left(\frac{\pi w}{W}\left(j + \frac{1}{2}\right)\right)$$

s. t. $h \in \{0, 1, \dots, H-1\}, w \in \{0, 1, \dots, W-1\}$

(1)

为了引入更多的信息解决通道注意力中的信息不充分问题,多光谱通道注意力模块将 GAP 推广到二维 DCT 表征频谱所有的特征信息,其结构如图 3 所示。其具体流程是,首先将输入 X 沿着通道维度划分为若干个部分,

$[X^0, X^1, \dots, X^{n-1}]$ 表示划分了 n 个部分,其中, $X^i \in R^{C' \times H \times W}, i \in \{0, 1, \dots, n-1\}, C' = \frac{C}{n}$ 。对于其中每个部分,分配相应的二维 DCT 频率分量,其结果可作为通道注意力的预处理结果,即:

$$Freq^i = 2DDCT^{u,v}(X^i) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} X_{i,h,w}^{u,v} B_{h,w}^{u,v}$$

s. t. $i \in \{0, 1, \dots, n-1\}$

(2)

其中, $[u, v]$ 为 X^i 的频率分量二维指标, $Freq \in R^C$ 为得到的多光谱向量,将上述各部分的频率分量串联起来得到:

$$Freq = compress(X) = cat([Freq^0, Freq^1, \dots, Freq^{n-1}])$$

(3)

整个多光谱通道注意力机制框架可以表示如下:

$$ms_att = sigmoid(fc(Freq))$$

(4)

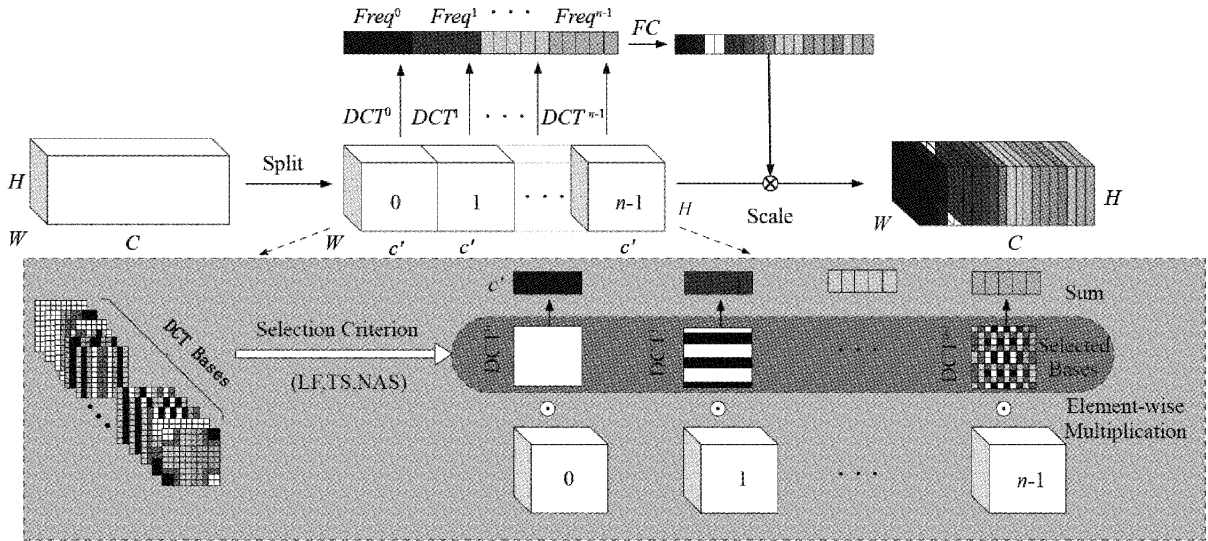


图 3 多光谱通道注意力模块网络结构

2.3 α -IoU 损失函数改进

CenterNet 采用了均方误差损失函数对边界框的宽和高进行预测,但均方误差损失函数独立计算检测框的中心点坐标和宽高等信息,造成损失值有较大的误差。在实际应用中目标检测常用交并比(intersection over union, IoU)代替均方误差损失函数来衡量预测框和真实框之间的重合度,但是 IoU 损失函数不能反映预测框和真实框之间的距离且无法精确的反映预测框和真实框的重合度大小,不利于边界框的精确定位。

因此在 CenterNet 中使用 α -IoU 损失函数^[20]来提高边界框预测的准确度, α -IoU 损失函数在 IoU 损失函数的基础上添加一个幂参数 α 进一步对边框回归损失函数优化,从而提升边界框回归和目标检测的准确性。 α -IoU 本质上是 $GIoU$ 、 $DIoU$ 、 $CIoU$ 等的统一幂化,其基础定义公式如下:

$$L_{\alpha-IoU} = \frac{1 - IoU^\alpha}{\alpha}, \alpha > 0$$

(5)

α -IoU 损失函数通过对公式中幂参数 α 取值进行调制,自适应地重新调整高和低 IoU 目标的梯度损失,可以很容易地改进检测器的效果,同时不会引入额外的参数,也不增加训练或者推理时间。本实验中选择训练效果最好的 $CIoU$ 损失函数改进为 α -IoU 损失函数,其中 $CIoU$ 和 α - $CIoU$ 损失函数公式如下:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b_{gt})}{c^2} + \beta\nu$$

(6)

$$L_{\alpha-CIoU} = 1 - IoU^\alpha + \frac{\rho^{2\alpha}(b, b_{gt})}{c^{2\alpha}} + (\beta\nu)^\alpha$$

(7)

式(6)、(7)中, b 表示预测框中心; b_{gt} 表示真实框中心; ρ 表示预测框与真实框中心点的欧式距离; c 表示能够包含预测框和真实框的最小闭包区域的面积; β 表示权重函数; ν 表示度量长宽的相似比。

$$\text{且: } \begin{cases} \nu = \frac{4}{\pi^2} (\arctan \frac{\tau w^{R'}}{h^{R'}} - \arctan \frac{\tau w}{h})^2 \\ \beta = \frac{\nu}{(1 - I_o U) + \nu} \end{cases} \quad (8)$$

3 实验结果与分析

本文实验采用的是 Windows 10 操作系统,实验平台使用的是基于 PyTorch1.2.0 的深度学习框架和 Python3.6 的代码解释器。实验环境配置如表 1 所示。

表 1 实验环境及配置

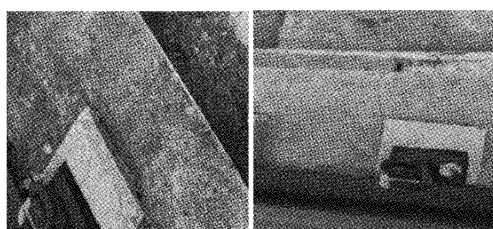
实验环境	实验配置
操作系统	Windows 10, 64 位
CPU	Intel i7-7800X
GPU	NVIDIA GeForce RTX 2080
内存	32 G
硬盘	256 G
显卡驱动	CUDA 10.0

3.1 轨道板裂缝数据集

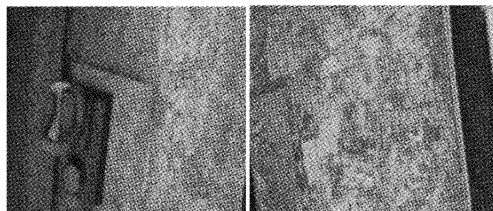
目前,关于高铁无砟轨道板表面裂缝分类研究缺乏相关的公共数据集。因此实验所用到的轨道板表面裂缝数据集图片是在轨道交通实训基地用相机固定拍摄。本文实验所采集到的轨道板表面裂缝数据集图片共约 3 000 多张,为提升实验检测效果,避免过拟合,随机采用翻转、缩放、裁剪、模糊、添加随机噪声数据增强等方法将数据集图片扩充到 6 200 张,并采用 labelimg 工具进行标注,然后将 VOC 数据集按照 9:1 的比例随机划分为训练集和测试集。其中训练集图片数量为 5 580 张,测试集图片数量为 620 张。该数据集的分类有两类,一类是 crack(代表条形裂缝),一类是 mapcrack(代表网状裂缝),裂缝数据集图片尺寸从 256×256 到 1 278×750 不等,数据集示例如图 4 所示。

3.2 模型评估指标

为从不同方面评价所改进方法对轨道板表面裂缝检测的有效性,本文实验主要采用均值平均精度(mean average precision, mAP)作为评价指标对检测模型的准确性进行评价。mAP 由所有类别的平均精度(average precision, AP)求均值得到,AP 是 PR 曲线、x 轴和 y 轴所



(a) 条形裂缝



(b) 网状裂缝

图 4 裂缝数据集

围成的近似面积。PR 曲线是由精确率(Precision)和召回率(Recall)组成的曲线。

3.3 消融实验

为了验证各个模块对传统的 CenterNet 算法的改进能力,对各模块进行评估,以判断每个模块的有效性。实验结果如表 2 所示,其中“√”表示添加相应模块,“×”表示未添加相应模块,数据中加粗字体表示各列最优结果。依次在原有算法基础上添加 ASPP 模块、MCA 模块和 α -IOU 损失函数,本文分别对各个模块进行了消融实验。可以发现,CenterNet 在轨道板裂缝数据集上的 $mAP = 80.75\%$,而改进后的 ASPP 模块的加入到 CenterNet 的主干网络中 mAP 提高了 1.15%,说明添加 ASPP 模块充分提取不同尺度的上下文信息使网络更加聚焦到裂缝上,检测性能得到提升;在添加 ASPP 模块的基础上再对 MCA 模块进行实验, mAP 又提高了 1.72%,说明 MCA 模块可以捕获图像丰富的输入特征信息;最终在 ASPP 模块和 MCA 模块基础上添加 α -IOU 损失函数,整体 mAP 相较之前模块提高了 0.5%,说明改进后的损失函数加大了惩罚措施有利于网络对裂缝进行精准定位,从而提高网络模型的性能。消融实验表明:相较传统的 CenterNet,本文改进算法通过引入更优的网络模块和边界框损失函数,在损失一点检测速度的情况下提高了模型的精度,改进后算法在 mAP 上提高了 3.37%,验证了各个模块融合的有效性。

表 2 改进算法消融实验

检测模型	ASPP	MCA	α -IoU	AP50/%		mAP/%	FPS
				crack	mapcrack		
CenterNet	×	×	×	87.46	74.05	80.75	44.55
CenterNet	√	×	×	88.85	74.96	81.90	40.75
CenterNet	√	√	×	90.05	77.19	83.62	38.61
CenterNet	√	√	√	89.38	78.86	84.12	37.67

3.4 实验效果对比

为了较直观地区分改进后的 CenterNet 算法与传统算法在两类裂缝的检测效果,从测试集中选取部分图片在相同实验环境下进行测试对比,其检测结果如图 5 所示。通

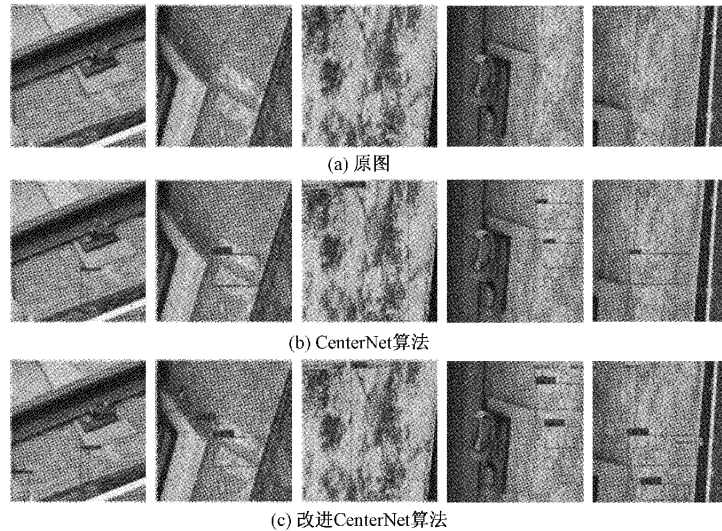


图 5 改进型 CenterNet 算法对比检测结果

3.5 不同算法检测能力对比

为了进一步验证算法的检测性能,本文将 CenterNet 算法与其他主流的深度学习算法进行对比,并选取均值平

均精度 mAP 和每秒处理帧数 FPS 作为不同检测算法的评价指标,得到实验结果如表 3 所示,其中数据中加粗字体为各列最优结果。

均精度 mAP 和每秒处理帧数 FPS 作为不同检测算法的评价指标,得到实验结果如表 3 所示,其中数据中加粗字体为各列最优结果。

表 3 实验结果对比

检测模型	主干网络	AP50/%		mAP/%	FPS
		crack	mapcrack		
Faster R-CNN	ResNet-50	85.21	73.62	79.42	15.86
SSD	VGG-16	77.94	75.31	76.63	48.64
RetinaNet	ResNet-50	68.66	75.13	71.89	29.15
YOLOv4	Darknet-53	89.89	64.34	77.12	32.73
CenterNet	ResNet-50	87.46	74.05	80.75	44.55
改进型 CenterNet	ResNet-50	89.38	78.86	84.12	37.67

根据表 3 实验结果对比可知,改进的算法模型较现阶段目标检测主流模型性能都有所提高。其中本文算法对轨道板表面裂缝检测精度为 84.12%,高于 Faster R-CNN、SSD、RetinaNet、YOLOv4 和 CenterNet 的检测精度。虽然检测速度不如 SSD 和 CenterNet,但是总的来看,本文算法在速度上仍具有巨大优势,达到 37.67 帧/s,可以满足实时性要求。通过实验证明,改进型 CenterNet 算法综合性能更好,可以更好地完成轨道板表面裂缝检测的任务要求。

4 结 论

本文为了解决传统的高铁无砟轨道板表面裂缝检测方法存在检测精度低,速度慢的问题,提出了一种基于改

进型 CenterNet 的轨道板表面裂缝检测算法。本算法采用 ASPP 扩大特征图的感受野,同时结合多光谱通道 MCA 更好的学习每个通道的权重,并使用 α -IoU 损失函数来提高边界框预测的准确度。实验表明,本文提出的改进型 CenterNet 的高铁无砟轨道板表面裂缝检测算法精度提升了 3.37%,该方法能够有效识别裂缝,适用于轨道板表面裂缝检测。但是随着网络模型复杂度的提高,检测速度有所下降。因此,我们在未来将继续探索如何减少网络模型参数量以及提升网络模型检测速度。

参考文献

- [1] 史云飞,何越磊,张志远,等.基于图像的 CRTS II 型轨道板面横向裂纹扩展特征分析[J].铁道科学与工程学报,2017,14(10):2059-2064.

- [2] 王登涛,李再伟,何越磊,等.基于热成像的高速铁路轨道板表面裂缝检测方法研究[J].铁道标准设计,2020,64(7):22-28.
- [3] ZHANG Y X, LI W, CHEN H, et al. CRTSII track slab crack detection based on improved YOLOv3 algorithm[C]. 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), IEEE, 2020: 1438-1441.
- [4] 李文举,何茂贤,张耀星,等.基于卷积神经网络和投票机制的轨道板裂缝检测[J].应用科学学报,2021,39(4):627-640.
- [5] YE W L, DENG S J, R J J. et al. Deep learning-based fast detection of apparent concrete crack in slab tracks with dilated convolution[J]. Construction and Building Materials, 2022, 32(9): 127-157.
- [6] REN S Q, HE K M, GIRSHICK R B, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [7] 王标,周雅兰,王永红.改进型 Faster R-CNN 网络在电子元件 LED 气泡缺陷检测中的应用[J].电子测量与仪器学报,2021,35(9):136-143.
- [8] He K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42(2): 386-397.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [10] 彭继慎,孙礼鑫,王凯,等.基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J].仪器仪表学报,2021,42(10):161-170.
- [11] 史梦圆,高俊钊.改进 YOLO v5 的建筑物破损检测算法研究[J].国外电子测量技术,2022,41(8):84-90.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [13] ZHOU X, WANG D, KRÄHENBÜHL P. Objects as points[J]. ArXiv Preprint, 2019, ArXiv:1904.07850.
- [14] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [15] YANG M K, YU K, ZHANG C, et al. DenseASPP for semantic segmentation in street scenes [C]. Proceedings of the 2018 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington: IEEE Computer Society, 2018: 3684-3692.
- [16] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2018: 7132-7141.
- [18] WANG Q, WU B, ZHU P, et al. ECA-net: Efficient channel attention for deep convolutional neural networks [C]. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Piscataway: IEEE, 2020: 11531-11539.
- [19] QIN Z, ZHANG P, WU F, et al. Fcanet: Frequency channel attention networks [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 783-792.
- [20] HE J, ERFANI S, MA X, et al. Alpha-IoU: A family of power intersection over union losses for bounding box regression [J]. Advances in Neural Information Processing Systems, 2021, 34: 20230-20242.

作者简介

吴铭权,硕士研究生,主要研究方向为计算机视觉、目标检测等。

罗晖,教授,主要研究方向为计算机视觉、图像处理等。
E-mail:1039758450@qq.com